HHS Public Access

Author manuscript

J Registry Manag. Author manuscript; available in PMC 2018 March 21.

Published in final edited form as:

J Registry Manag. 2012 ; 39(3): 101–106.

Evaluating Factors Associated with Unknown SEER Summary Stage 2000 Derived from Collaborative Stage

Mei-Chin Hsieh, MSPH, CTR[1], **Qingzhao Yu, PhD**[2], **Xiao-Cheng Wu, MD, MPH, CTR**[1], **Brad Wohler, MS**[3], **Ying Fan, MS**[2], **Baozhen Qiao, PhD**[4], **Ahmedin Jemal, DVM, PhD**[5], and **Umed A. Ajani, MBBS, MPH**[6]

[1]Louisiana Tumor Registry and Epidemiology Program, School of Public Health, Louisiana State University Health Sciences Center, New Orleans, LA

[2]Biostatistics Program, School of Public Health, Louisiana State University Health Sciences Center, New Orleans, LA

[3]Florida Cancer Data System, Miami, FL

[4]New York State Cancer Registry, New York State Department of Health, Albany, NY

[5]American Cancer Society, Atlanta, GA

[6]Division of Cancer Prevention and Control, Centers for Disease Control and Prevention, Atlanta, GA

Abstract

Background—Information on cancer stage is critical for guiding treatment and assessing disease prognosis. The Percentage of unknown staged cancer cases varies considerably across state cancer registries; factors contributing to the variations in unknown stage have not been reported in the literature before. The purpose of this study was to examine whether these variations are influenced by demographic and clinical factors as well as type of reporting facilities.

Methods—Invasive colorectal, lung, female breast, and prostate cancers diagnosed in 2004–2007 and staged (derived Summary Stage 2000) according to Collaborative Stage Version 1 were obtained from the North American Association of Central Cancer Registries (NAACCR); 47 population-based cancer registries in the United States were included. Relative importance analysis was used to identify variables that were relatively important in predicting unknown stage. We used multiple linear regression to evaluate factors associated with percentage of unknown stage by cancer site using state central cancer registries as analytic units; potential outlier registries with high percentage of unstaged cases were identified using boxplots and standardized residuals.

Results—Overall, lung cancer had the highest percentage of unknown stage (8.3%) and prostate cancer had the largest variation of unknown stage among registries (0.6%–18.1%). The

Corresponding author: Mei-Chin Hsieh, MSPH, CTR, Louisiana Tumor Registry, School of Public Health, Louisiana State University Health Sciences Center, 2020 Gravier St, 3rd Floor, New Orleans, LA 70112, Phone: (504) 568-5850, Fax: (504) 568-5800, mhsieh@lsuhsc.edu.

Disclaimers:

The authors declare no financial conflicts of interest in this work. The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

percentages of neoplasm Not Otherwise Specified (NOS) histology, non-microscopically confirmation, and non-hospital reporting source were positively associated ($p<0.05$) with percentage of unknown stage for all studied cancer sites before adjustment. Variables that retained a positive association with unknown stage after adjusted for demographic variables, clinical variables, year of diagnosis, and type of reporting source were black race, metropolitan area < 1 million population, histologies of neoplasm NOS or epithelial neoplasm NOS, diagnosis year 2005, and non-hospital reporting source for colorectal cancer; metropolitan area < 1 million population, neoplasm NOS histology, and non-hospital reporting source for female breast; and diagnosis year 2005 and non-hospital reporting source for prostate. After adjustment, none of the predictors were significant for lung cancer. We observed one potential outlier registry each for colorectal, lung, and female cancers.

Conclusions—Factors associated with unknown stage differ by cancer site; however, type of reporting source is an important predictor of unknown stage for all cancers except lung after adjustment. Cancer registries with high percentage of unknown stage should be made aware of their data quality issue(s). As a result, these registries can investigate those factors and provide training to registrars to improve their cancer data quality.

Keywords

Cancer; Summary Stage; Collaborative Stage; Cancer Registry

Introduction

Information on cancer staging is essential for assessing the effectiveness of early detection, intervention, planning treatment, and predicting outcome. Formed in 1998, the Collaborative Staging (CS) Task Force,[1] now known as the CS Governance Committee, developed the CS Data Collection System for use with cases diagnosed in 2004 and after to resolve the issue of discrepancies in staging guidelines and to provide a higher degree of compatibility among the three major cancer staging systems: American Joint Committee on Cancer (AJCC) staging,[2] Surveillance, Epidemiology, and End Result (SEER) Summary Stage,[3] and SEER Extent of Disease (EOD).[3] An additional advantage of using CS is the reduction in unknown stage rates. A study using data in the CDC's National Program of Cancer Registries (NPCR) Cancer Surveillance System (CSS) observed a decreased percentage of unknown SEER Summary Stage 2000 (SS2000) from 2001–2003 to 2004–2005.[4]

The percentage of unknown stage is a strong quality indicator of stage data as well as the quality of abstraction and availability of source data. The Data Assessment Workgroup of the North American Association of Central Cancer Registries (NAACCR) Data Use and Research Committee (DURC) found that the percentage of unknown stage cases varies substantially by cancer registry and cancer site.[5] The common approach in handling unknown stage cancer cases in research is to exclude them from the statistical analysis; however, this could create biased results.[6] To help registries to implement appropriate strategies to improve quality of stage data, it is important to identify and quantify factors associated with unknown stage. Unfortunately these studies are lacking in the literature. Previous studies on cancer stage primarily focused on finding factors associated with advanced stage only.[7–11] The purposes of this study were to identify factors associated with

unknown stage cases at central registry level for colorectal, lung, female breast, and prostate cancer as well as to identify registries with unusual high percentages of unstaged cases.

Data and Methods

Data Source

Data from 47 United States (US) population-based cancer registries were obtained from the NAACCR Cancer in North America (CINA) Analytic file 1995–2007 data which included 48 US state cancer registries (Maryland and Nevada were not included in this analytic file). Minnesota was excluded because the Rural/Urban 2003 Continuum code was not available. We included invasive colorectal (ICD-O-3[12] Topography: C18.0–C18.9, C19.9, C20.9), lung (C34.1–C34.9), female breast (C50.0–C50.9), and prostate cancer (C61.9) diagnosed in the years 2004 to 2007. Autopsy or death-certificate-only cases, lymphomas originating in the sites of interest, and cases with unknown Rural/Urban information were excluded from the analysis.

The outcome of interest was the percentage of cases with unknown SEER Summary Stage 2000 (SS2000) derived from Collaborative Stage Version 1 (CSv1) at the central registry level. Independent variables were demographic (i.e., race, sex, age, and rural/urban residence) and clinical (i.e., histology, grade, and diagnostic confirmation) data, as well as year of diagnosis and type of reporting source. Residence regions were defined based on the 2003 Rural/Urban Continuum codes[12]: county metropolitan areas with populations of one million or more, county metropolitan areas with populations of less than one million, and non-metropolitan areas. Histologies were categorized as: neoplasms, Not Otherwise Specified (NOS) (ICD-O-3[13] morphology 8000–8005), epithelial neoplasms NOS (8010–8046), adenocarcinoma NOS (8140), and specified histologies. The type of reporting source was grouped into two categories: hospital and non-hospital facilities. Beginning with cases diagnosed in 2006, two new categories, radiation treatment centers/medical oncology centers (hospital-affiliated or independent) and other hospital outpatient units/surgery centers, were introduced as types of reporting sources in NAACCR records.[14–15] The hospital group included: hospital inpatient, radiation treatment centers/medical oncology centers, and hospital outpatient units/surgery centers. The non-hospital facilities included: physician’s office, nursing home/hospice, and laboratory only.

Statistical Analysis

The relative importance of variables[16] was adopted to describe the contribution of each variable in explaining the variance of the outcome. In detail, independent variables were sequentially added to the linear regression models, from which the increase in R square was recorded for each variable. We permuted the order of variables to be added in the linear models. The relative importance of the variables is defined as the average increase in R square from the permutation for each variable. Variable with high relative importance index indicates that the predictor is relatively important in predicting the percentage of unknown stage, which was treated as a continuous variable ranging variously by cancer site.

Multiple linear regression was used to evaluate the relationship between the predictors and the percentage of cases with unknown stage for selected cancer sites and the regression models were run separately for each cancer site. The covariate was the percentage of patients falling into each particular subgroup of a predictor at central registry level. Therefore a predictor with k categories forms k-1 covariates. To identify registries with unusually high percentages of unknown stage, we used the upper whisker of the boxplots of the unadjusted percentages and the standardized residuals generated from the multiple linear regression model that adjusted for important factors. The upper whisker of the boxplots is defined as the largest data point of unknown stage percentage within the boundary of following formula: Third quartile + 1.5× (Third quartile- First quartile). Any registries with standardized residuals of unknown stage percentage higher than the upper bound of 95% conference interval (CI) were considered to have unusually high percentage of unstaged cases. The statistical significant level was set at 0.05. All analyses were carried out using SAS® version 9.2 (SAS Institute, Cary, NC).

Results

Lung cancer had the highest percentage of unknown derived SS2000 (8.3%), followed by colorectal cancer (7.2%) and prostate cancer (6.6%); female breast cancer had the lowest percentage of unstaged cases (3.5%) for all cases combined (Table 1). At the registry level, the ranges of percentages of unknown stage varied substantially for all studied cancer sites, particularly for prostate cancer with a range from 0.6%–18.1% (Table 1). The three most important predictors based on relative importance analysis for colorectal cancer were histology type, type of reporting source, and tumor grade; for lung cancer, type of reporting source, diagnostic confirmation, and age; for female breast cancer, type of reporting source, histology type, and diagnostic confirmation; and for prostate cancer, tumor grade, type of reporting source, and urban/rural area (Figure 1).

Colorectal Cancer

Registries with higher percentages of patients residing in non-metropolitan areas or with higher percentages of moderately differentiated tumor grade had lower percentages of unknown stage (Table 2), whereas registries with higher percentages of neoplasms NOS histology, higher percentages of unknown diagnostic confirmation, or higher percentages of non-hospital cases had higher percentages of unknown stage in the univariate analysis. After adjustment for all predictors, the percentages of cases with unknown stage differed significantly by race, rural/urban residence, histology type, diagnostic confirmation, diagnosis year and type of reporting source. The difference was particularly strong for histology type. Compared with specified histology, a 1% increase of neoplasms NOS or a 1% increase of epithelial neoplasms NOS was related to an average of 2.7% or 3.7% increase in unstaged rate after controlling for other predictors (Table 2).

Lung Cancer

When other predictors were not controlled, older age groups were related to lower percentages of unknown stage than those aged less than 50 years old (Table 2), and registries with higher percentages of neoplasms NOS or unknown microscopic confirmation or non-

hospital reporting sources were associated with higher percentages of unknown stage. However, after adjusting for other predictors, no predictors were significantly associated with unknown stage.

Female Breast Cancer

The percentages of cases with histologies of neoplasm NOS or unknown diagnostic confirmation or non-hospital reporting source had positive associations with unknown stage in univariate analysis. This is most striking for unknown diagnostic confirmation. Compared with microscopically confirmed cases, a 1% increase of unknown diagnostic confirmation rate related to about an average of 6% increase in the percentage of unknown stage (Table 2). After adjusting for other predictors, the percentage of patients residing in metropolitan areas with populations less than one million, the percentage coded to neoplasms NOS, and the percentage abstracted at non-hospital sources were positively associated with unknown stage. Compared to specified histology, a 1% increase of neoplasms NOS was related to an average of 3.2% increase in unstaged rate after other covariates were controlled.

Prostate Cancer

Although tumor grade yielded the highest importance index in the variable relative importance test (Figure 1), no statistically significant differences resulted when the percentages of well-differentiated cases were compared with other grade groups in the model that was not adjusting for other predictors (Table 2). However, the F-test showed that tumor grade was statistically significantly associated with percentage of unknown stage (p-value=0.014). We further examined the association using unknown grade as baseline and found that grades of moderately and poorly differentiated had negative associations with the percentage of unknown stage (coefficients −0.731 and −0.579, respectively). After adjustment, registries with higher percentages of cases diagnosed in 2005 or higher percentage of cases reported by non-hospital reporting sources were more likely to have a higher percentage of unstaged prostate cancer cases.

Identifying registries with high unstaged rate

Before adjusting for all important predictors, we identified two registries as outliers with high unstaged rates (above upper whisker) for colorectal cancer, one for lung, six for breast and six for prostate cancer (Table 1). After adjustment, we found that only two registries had unusually high percentages of unknown stage (outside the upper bound of 95% CI), one for both colorectal and female breast cancers and one for lung cancer.

Discussion

The percentage of unknown stage varied by registry and by cancer site. Overall, female breast cancer had the lowest percentage of unknown stage cases among the four sites. Neoplasm NOS, unknown diagnostic confirmation, and non-hospital reporting source were positively associated with high percentages of unknown stage for all four cancer sites. Although race has been associated with advanced stage for colorectal, female breast, and prostate cancer,[7–11] it was significantly associated with unknown stage only for colorectal cancer after controlling for other predictors in our study. Diagnosis year alone was

associated with unknown stage only in 2005 for colorectal and prostate cancer after adjustment. Colorectal cancer without microscopic diagnostic confirmation had a lower percentage of unstaged cases in the linear regression model (a 1% increase in the percentage of cases lacking diagnostic confirmation rate reduced the rate of unknown stage by about 4% after controlling for other covariates). The reasons for this unexpected result could be that diagnostic confirmation interacts with other predictors and/or does not fit in an ordinary linear regression model. Some predictors with high relative importance indexes were not significantly associated with unknown stage after adjustment because they are not significantly different from other predictors in terms of relative importance. Although breast and prostate cancers had higher numbers of registries with high percentage of unstaged cases; after adjustment, only one registry retained a high percentage of unknown stage, and that was for breast cancer only.

We did not observe a consistent decrease in the percentage of unknown stage over time as anticipated, except for lung cancer. However, the decrease in the unknown stage percentage for lung cancer diagnosed in 2005, 2006, or 2007 was not significantly different from cases diagnosed in 2004. For colorectal, female breast, and prostate cancers the changes in the percentage of unknown stage varied by diagnosis year. The failure to observe a declining trend in unknown stage could be due to a short observation time period and/or improperly interpreted or unclear CS coding guidelines. This study included only four years of data, from diagnosis year 2004 to 2007. During this time period, several updates and changes occurred in the Collaborative Staging coding schema. Although several training and education webinars and workshops were conducted before and after the CS Data Collection System was implemented, some of the coding guidelines might not have clearly explained the use of coding rules such as code none versus unknown.

As expected, state central registries with a higher percentage of cases diagnosed in non-hospital facilities yielded a higher percentage of unknown stage than registries having higher percentage of cases diagnosed in hospitals before and after adjustment, except for lung cancer after adjustment. This could reflect lack of information in non-hospital charts and underscore the need for proper and complete data collection from non-hospital facilities.

The multiple imputation method has been commonly used in dealing with missing values.[17] Eisemann et al. showed it is a useful technique for imputing missing values for tumor stage except for cancer sites with high percentages of missing stage information.[18] Therefore, it is critical that cancer registries reduce the unknown stage percentage as well as unknown percentages for other important variables to generate more valid results when using registry data.

This study has several strengths. First, it included 47 of the 50 US state cancer registries; therefore, the results are generalizable of the whole US. Second, data were collected and coded uniformly based on CS and NAACCR rules, ensuring that coding schemes for variables used in this study were the same across registries. Third, we used central registry level data as an analysis unit, instead of the individual cancer case level; this can identify issues occurring at individual registries and can also identify potential outlier registries.

Despite these strengths, there are a few limitations to this study. The registries' experience of coding SEER Extent of Disease (EOD)[3] was not taken into account in the analysis. Coding rules for tumor size, tumor extension and lymph node involvement in SEER EOD were adapted and modified in the CS system. Registries that collected SEER EOD for cases diagnosed before 2004 may have been more likely to accept the concept of coding CS data items. Another limitation is that class of case (NAACCR item 160)[14] information was not available in our study. Non-analytic cases (Class of case codes 3–9),[14] which include cases not diagnosed and/or treated at reporting facility, provide very limited information on CS-related information; therefore, registries with high percentages of non-analytic cases may increase the unstaged percentage.

In summary, our study has shown that for colorectal, lung, female breast, and prostate cancers the important predictors of unknown percentage of derived SEER Summary Stage 2000 at the univariate level include histologic type, diagnostic confirmation, and type of reporting source. The relative importance of these factors, however, varies by site. For cancer registries having a higher percentage of unknown stage, further investigation of factors that caused higher unknown stage percentage is needed. Population-based cancer registry data are a valuable source for cancer research; any effort to reduce the percentage of unstaged cases will improve the quality of cancer data and produce more reliable cancer research results. A future study on changes in unknown staged rates over time is needed as more data become available.

Acknowledgments

The authors would like to thank the members of the NAACCR Data Assessment Workgroup, particularly Patricia A. Andrews, Maria J. Schymura, Hannah K. Weir, Missy Jamison, Bin Huang, and Xiangrong Li, for their general support.

References

1. Collaborative Stage Data Collection System. [Accessed September 25, 2011] Available at http://www.cancerstaging.org/cstage/about.html
2. [Accessed September 25, 2011] American Joint Committee on Cancer (AJCC) staging. Available at http://www.cancerstaging.org
3. [Accessed September 25, 2011] Surveillance Epidemiology and End Results: Historical Staging and Coding Manuals. Available at http://seer.cancer.gov/tools/codingmanuals/historical.html
4. Wilson, RJ., He, Q., Thomas, CC., et al. Evaluation of summary stage data quality in the CDC-NPCR Cancer Surveillance System. Paper presented at: Annual Meeting of North American Association of Central Cancer Registries; June 16, 2009; San Diego, CA.
5. Yu, Q., Wu, XC., Hsieh, MC., et al. Cancer data quality control by proportion of unknown stage. Paper presented at: Annual Meeting of North American Association of Central Cancer Registries; Louisville KY. June 2011;
6. Greenland S, Finkle WD. A critical look at methods for handling missing covariates in epidemiologic regression analysis. Am J Epidemiol. 1995; 142:1255–1264. [PubMed: 7503045]
7. Schwartz KL, Crossley-May H, Vigneau FD, et al. Race, socioeconomic status and stage at diagnosis for five common malignancies. Cancer Causes and Control. 2003; 14:761–766. [PubMed: 14674740]
8. Parikh-Patel A, Bates JH, Campleman S. Colorectal cancer stage at diagnosis by socioeconomic and urban/rural status in California, 1988–2000. Cancer (suppl). 2006; 107(5):1189–1195.

9. Hahn KM, Bondy ML, Selvan M, et al. Factors associated with advanced disease stage at diagnosis in a population-based study of patient with newly diagnosed breast cancer. Am J Epidemiol. 2007; 166:1035–1044. [PubMed: 17690220]

10. Li CI, Malone KE, Daling JR. Differences in breast cancer stage, treatment, and survival by race and ethnicity. Arch Intern Med. 2003; 163:49–56. [PubMed: 12523916]

11. Hoffman RM, Gilliland FD, Eley JW, et al. Racial and ethnic differences in advanced-stage prostate cancer: the prostate cancer outcomes study. J Natl Cancer Inst. 2001; 93:388–395. [PubMed: 11238701]

12. [Accessed September 25, 2011] Rural/Urban Continuum Codes. 2003. Available at http://www.ers.usda.gov/data/ruralurbancontinuumcodes/2003/

13. Fritz, A.Percy, C.Jack, A.Shanmugaratnam, K.Sobin, L., Parkin, DM., editors. ICD-O-3: International classification of diseases for oncology. 3. Geneva: World Health Organization; 2000.

14. [Accessed October 10, 2011] NAACCR Data Standards and Dictionary (Vol. II) Version 11. Available at http://www.naaccr.org/LinkClick.aspx?fileticket=AL3xNvPVuv0%3d&tabid=133&mid=473

15. [Accessed December 20, 2011] NAACCR Data Standards and Dictionary (Vol. II) Version 10. Available at http://www.naaccr.org/LinkClick.aspx?fileticket=AfNCNY4EPzc%3d&tabid=133&mid=473

16. Lindeman, RH., Merenda, PF., Gold, RZ. Introduction to Bivariate and Multivariate Analysis. Scott, Foresman; Glenview, IL: 1980. p. 119

17. Rubin DB. Multiple imputation after 18+ years. Journal of the American Statistical Association. 1996; 91:473–489.

18. Eisemann N, Waldmann A, Katalinic A. Imputation of missing values of tumour stage in population-based cancer registration. BMC Medical Research Methodology. 2011; 11:129. [PubMed: 21929796]

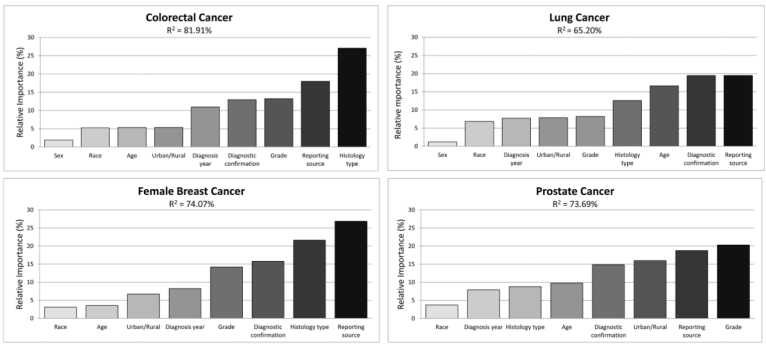

Figure 1.
Variable relative Importance by cancer site

Table 1

Measures of central tendency and dispersion on unknown derived summary stage 2000 by cancer site from 47 United States population-based cancer registries 2004–2007

	Colorectal	Lung	Female Breast	Prostate
Mean for all cases combined	7.17% (N=552,242)	8.25% (N=761,739)	3.51% (N=746,887)	6.59% (N=781,923)
Based on individual registry				
Mean	6.40%	7.30%	3.11%	5.34%
25–75% Percentile (IQR[1])	4.61%–7.27%	4.95%–9.03%	1.56%–3.27%	2.59%–6.44%
Minimum	2.40%	2.44%	0.98%	0.61%
Median	5.48%	6.55%	2.29%	3.78%
Maximum	18.80%	18.68%	13.68%	18.06%
[2] Upper Whisker	11.25%	15.16%	5.83%	12.23%
[3] Lower Whisker	2.40%	2.44%	0.98%	0.61%
Number of registry outside of upper whisker	2	1	6	6

[1] IQR: Interquartile range

[2] Upper Whisker: extends to largest data point within the boundary of Q3 + 1.5* (Q3−Q1)

[3] Lower Whisker: extends to smallest data point within the boundary of Q1 − 1.5* (Q3−Q1)

J Registry Manag. Author manuscript; available in PMC 2018 March 21.

Table 2

Coefficients of unadjusted and reduced model by cancer site: 47 United States population-based cancer registries 2004–2007

	Colorectal		Lung		Female Breast		Prostate	
Variable[§]	**Unadjusted**	**Adjusted**	**Unadjusted**	**Adjusted**	**Unadjusted**	**Adjusted**	**Unadjusted**	**Adjusted**
Sex								
Female	Referent	Referent	Referent	Referent	~	~	~	~
Male	−0.348	−0.216	−0.348	−0.123	~	~	~	~
Race								
White	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
Black	0.038	0.131*	0.027	−0.058	0.037	0.039	0.070	0.072
Other races	−0.028	0.032	−0.052	−0.114	−0.009	0.031	0.030	−0.042
Age								
≤49	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
50–64	−0.565	−1.866	−3.214*	−2.107	−0.390	−0.608	−2.575	3.330
65–74	−0.321	−0.910	−1.901*	−1.250	0.001	−0.063	−2.719	1.744
75+	−0.320	−1.027	−2.347*	−1.895	−0.360	−0.292	−2.367	2.845
Rural/urban residence								
Metro area ≥ 1 million pop	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
Metro area < 1 million pop	0.003	0.108*	−0.038	−0.039	0.002	0.067*	−0.073	−0.017
Non-metro area	−0.040*	−0.009	−0.031	−0.017	−0.024	−0.020	−0.071*	−0.028
Histology type								
Specified histologies	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
Neoplasms, NOS†	2.486*	2.718*	0.719*	0.925	2.629*	3.168*	2.031*	−1.670
Epithelial neoplasms, NOS†	1.227	3.653*	0.159	0.505	0.364	0.395	−0.475	1.033
Adenocarcinoma, NOS†	0.197	0.102	0.173	0.684	0.636	−0.245	0.007	−0.235
Grade								
Well differentiated	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
Moderately differentiated	−0.507*	0.311	−0.836	−0.095	−0.277	−0.291	0.320	−0.624
Poorly differentiated	−0.121	−0.129	−0.090	0.475	0.012	−0.061	0.327	−0.680
Undifferentiated	−0.939	1.290	−0.249	0.228	−0.761	−0.434	0.361	−0.517

Variable[§]	Colorectal		Lung		Female Breast		Prostate	
	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted	Unadjusted	Adjusted
Unknown	0.066	0.317	−0.300	0.152	0.184	−0.247	1.080	0.410
Diagnostic confirmation								
Yes	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
No	−1.132	−3.863*	−0.146	−0.397	1.282	−1.197	−2.541	−2.357
unknown	2.910*	−1.549	1.076*	0.218	5.999*	−0.553	3.114*	2.112
Diagnosis year								
2004	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
2005	1.491	1.497*	−0.377	−0.218	0.482	0.102	−0.331	1.541*
2006	−0.508	−0.088	−0.626	−0.665	0.446	0.675	−0.865	−0.152
2007	−0.766	−0.0004	−1.538	−0.741	−0.683	−0.484	0.426	0.391
Type of reporting source								
Hospital	Referent	Referent	Referent	Referent	Referent	Referent	Referent	Referent
Non-Hospital	0.422*	0.356*	0.380*	0.238	0.302*	0.223*	0.240*	0.293*

[§] Predictor of k categories forms k-1 covariates and covariate is the proportion of patients falling in the subgroups of each predictor at each registry

~ Not applicable

† NOS indicates Not otherwise specificed

* Statistical significance at p-value ≤0.05